

A DETECTION OF WATER IN THE TRANSMISSION SPECTRUM OF THE HOT JUPITER WASP-12b AND IMPLICATIONS FOR ITS ATMOSPHERIC COMPOSITION

LAURA KREIDBERG^{1,8}, MICHAEL R. LINE², JACOB L. BEAN^{1,9}, KEVIN B. STEVENSON^{1,10}, JEAN-MICHEL DÉSERT³, NIKKU MADHUSUDHAN⁴, JONATHAN J. FORTNEY², JOANNA K. BARSTOW⁵, GREGORY W. HENRY⁶, MICHAEL H. WILLIAMSON⁶, & ADAM P. SHOWMAN⁷

Draft version October 20, 2015

ABSTRACT

Detailed characterization of exoplanets has begun to yield measurements of their atmospheric properties that constrain the planets' origins and evolution. For example, past observations of the dayside emission spectrum of the hot Jupiter WASP-12b indicated that its atmosphere has a high carbon-to-oxygen ratio ($C/O > 1$), suggesting it had a different formation pathway than is commonly assumed for giant planets. Here we report a precise near-infrared transmission spectrum for WASP-12b based on six transit observations with the *Hubble Space Telescope*/Wide Field Camera 3. We bin the data in 13 spectrophotometric light curves from $0.84 - 1.67 \mu\text{m}$ and measure the transit depths to a median precision of 51 ppm. We retrieve the atmospheric properties using the transmission spectrum and find strong evidence for water absorption (7σ confidence). This detection marks the first high-confidence, spectroscopic identification of a molecule in the atmosphere of WASP-12b. The retrieved 1σ water volume mixing ratio is between $10^{-5} - 10^{-2}$, which is consistent with $C/O > 1$ to within 2σ . However, we also introduce a new retrieval parameterization that fits for C/O and metallicity under the assumption of chemical equilibrium. With this approach, we constrain C/O to $0.5^{+0.2}_{-0.3}$ at 1σ and rule out a carbon-rich atmosphere composition ($C/O > 1$) at $> 3\sigma$ confidence. Further observations and modeling of the planet's global thermal structure and dynamics would aid in resolving the tension between our inferred C/O and previous constraints. Our findings highlight the importance of obtaining high-precision data with multiple observing techniques in order to obtain robust constraints on the chemistry and physics of exoplanet atmospheres.

Subject headings: planets and satellites: atmospheres — planets and satellites: composition — planets and satellites: individual: WASP-12b

1. INTRODUCTION

The chemical composition of a planetary atmosphere provides a rich record of the planet's formation conditions and evolutionary history. Measurements of the composition can constrain the planet's formation mechanism, its formation location in the protoplanetary disk, the surface density and composition of planetesimals at the formation site, the relative accretion rates of gas and solids, and possible migration pathways (e.g. Atreya et al. 1999; Owen et al. 1999; Gautier et al. 2001; Atreya et al. 2003; Hersant et al. 2004; Lodders 2004; Dodson-Robinson et al. 2009; Öberg et al. 2011; Madhusudhan et al. 2011a, 2014a).

Because there are many factors that influence atmo-

spheric chemical composition, a large sample size of planets is required to develop a comprehensive theory of giant planet formation. Fortunately, the sample of known extrasolar giant planets is large and growing, and recent observations of these planets have begun to yield basic constraints on their atmospheric chemistry. These include inferences of the carbon-to-oxygen ratio (C/O) and absolute water abundance, which in some cases rival our knowledge of those quantities for the Solar System planets (e.g. Madhusudhan et al. 2011b; Konopacky et al. 2013; Line et al. 2014; Kreidberg et al. 2014a; Brogi et al. 2014; Madhusudhan et al. 2014b).

One of the best studied exoplanet atmospheres is that of the transiting hot Jupiter WASP-12b. This planet is a $1.4 M_{\text{Jup}}$, $1.8 R_{\text{Jup}}$ gas giant orbiting a late-F host star with a period of just 1.1 days (Hebb et al. 2009). The brightness of the host star ($H=10.2$), the planet's high equilibrium temperature (2500 K), and the planet's large size make WASP-12b a favorable target for atmosphere characterization. The system has been observed extensively from the ground and space to measure the planet's transmission spectrum, dayside emission spectrum, and thermal phase variation (López-Morales et al. 2010; Fossati et al. 2010; Campo et al. 2011; Croll et al. 2011; Cowan et al. 2012; Crossfield et al. 2012; Haswell et al. 2012; Zhao et al. 2012; Föhring et al. 2013; Sing et al. 2013; Swain et al. 2013; Mandell et al. 2013; Copperwheat et al. 2013; Stevenson et al. 2014c,b; Burton et al. 2015; Nichols et al. 2015; Croll et al. 2015).

E-mail: laura.kreidberg@uchicago.edu

¹ Department of Astronomy and Astrophysics, University of Chicago, 5640 S. Ellis Ave, Chicago, IL 60637, USA

² Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA

³ CASA, Department of Astrophysical & Planetary Sciences, University of Colorado, 389-UCB, Boulder, CO 80309, USA

⁴ Institute for Astronomy, University of Cambridge, Cambridge CB3 OHA, UK

⁵ Department of Physics, University of Oxford, Oxford OX1 3RH, UK

⁶ Center of Excellence in Information Systems, Tennessee State University, Nashville, TN 37209, USA

⁷ Department of Planetary Sciences and Lunar and Planetary Laboratory, The University of Arizona, Tucson, AZ 85721, USA

⁸ National Science Foundation Graduate Research Fellow

⁹ Packard Fellow

¹⁰ Sagan Fellow

An intriguing possibility has emerged from these observations of WASP-12b, which is that the planet has a carbon-rich atmospheric composition ($C/O > 1$, compared to the solar value of 0.55; Asplund et al. 2009). The high C/O interpretation was first suggested by Madhusudhan et al. (2011b) as the best explanation for the planet’s dayside emission spectrum. This result was contested by subsequent work (Crossfield et al. 2012; Line et al. 2014); however, the most recent comprehensive analysis of the dayside spectrum reaffirmed the inference of high C/O (Stevenson et al. 2014b).

The claim of high C/O in WASP-12b’s atmosphere has motivated substantial theoretical and observational work. This includes attempts to measure C/O for additional planets (e.g. Brogi et al. 2014; Line et al. 2014), studies of the effect of C/O on atmospheric chemistry (Madhusudhan et al. 2011a; Madhusudhan 2012; Kopparapu et al. 2012; Moses et al. 2013; Venot et al. 2015), inferences of C/O in exoplanet host stars (Teske et al. 2013, 2014), and predictions of C/O from planet formation theory (Madhusudhan et al. 2011a; Öberg et al. 2011; Madhusudhan et al. 2014a; Ali-Dib et al. 2014).

We note, however, that the evidence for a carbon-rich atmosphere on WASP-12b is not definitive. The high C/O inference is based primarily on the photometric secondary eclipse depth from *Spitzer* at $4.5 \mu\text{m}$. The low brightness temperature for the planet in this bandpass is best explained by absorption from carbon-bearing species (either CO, HCN, or C_2H_2 ; Madhusudhan et al. 2011b; Kopparapu et al. 2012; Moses et al. 2013; Stevenson et al. 2014c), but photometry alone cannot uniquely identify which molecule is the dominant absorber. No other molecular features have been confidently identified in the emission spectrum, and measurements of the planet’s transmission spectrum have yielded even less conclusive constraints on the C/O (Sing et al. 2013; Mandell et al. 2013; Swain et al. 2013; Stevenson et al. 2014c; Madhusudhan et al. 2014b). High-precision spectroscopy is thus needed to obtain unambiguous determination of the atmospheric composition of this important planet.

In this work, we report a new, precise measurement of WASP-12b’s transmission spectrum over the wavelength range 0.84 to $1.67 \mu\text{m}$. The outline of the paper is as follows. In § 2 we present the observations and data reduction. The light curves fits and measurement of the transmission spectrum are outlined in § 3. We compare our results to previous measurements in § 4. We describe a retrieval of the planet’s atmospheric properties based on this transmission spectrum in § 5 and discuss implications for the chemical composition in § 6. We conclude in § 7.

2. OBSERVATIONS AND DATA REDUCTION

2.1. *Observations*

We obtained time series spectroscopy during six transits of WASP-12b between UT 1 January and 4 March 2014 using the Wide Field Camera 3 (WFC3) IR detector on the *Hubble Space Telescope* (*HST*) as part of *HST* GO Program 13467. Three of the transit observations used the G102 grism, which provides spectral coverage from 0.82 – $1.12 \mu\text{m}$, and the other three used the G141 grism, which spans the range 1.12 – $1.65 \mu\text{m}$. The G141 grism has been widely used for exoplanet transit spec-

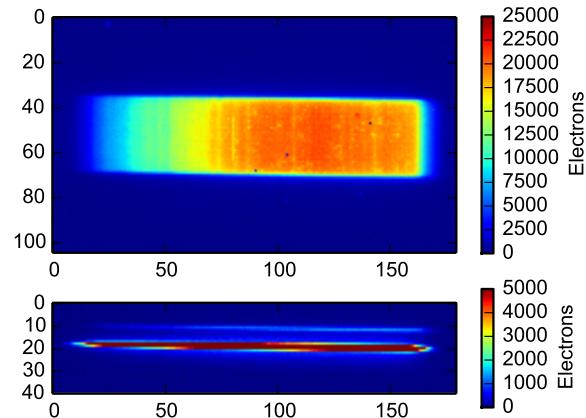


FIG. 1.— Raw *HST*/WFC3 images taken with the G102 grism. Spatial scan and staring mode data are shown in the top and bottom panels, respectively. The images are cutouts from a 256x256 pixel subarray. The spectrum of WASP-12A’s binary companion is visible in the bottom panel near row 10.

troscopy, but these observations are the first to use the G102 grism for this purpose. The additional wavelength coverage from the G102 grism provides access to features from more molecules and absorption bands, giving us greater leverage in constraining the atmospheric composition of the planet. Each transit observation (called a visit) consisted of five consecutive, 96-minute *HST* orbits. WASP-12 was visible during approximately 45 minutes per orbit and occulted by the Earth for the remainder of the time. We took a direct image of the target with the F126N narrow-band filter at the beginning of each orbit for wavelength calibration. Example staring mode and spatial scan images are shown in Figure 1.

The spectroscopic data were obtained in spatial scan mode with the 256x256 subarray, using the SPARS10, NSAMP=16 readout pattern, which has an exposure time of 103.1 s. The scan rates were 0.04 and 0.05 arc-sec/second for the G102 and G141 grisms, respectively. The spectra extend roughly 40 pixels in the spatial direction, with peak per-pixel counts below 25,000 electrons for both grisms. We alternated between forward and reverse scanning along the detector to decrease instrumental overhead time. This setup yielded 19 exposures per orbit and a duty cycle of 74%.

In addition to the spatial scan data, we also obtained 10 staring mode spectra in each grism during the first orbit of the first visit. WASP-12 is a triple star system: WASP-12 A hosts the planet, and WASP-12 BC is an M-dwarf binary separated from WASP-12 A by about $1''$ (Bergfors et al. 2013; Bechter et al. 2014). We use these staring mode data to resolve the spectrum of WASP-12 A from WASP-12 BC, enabling us to correct for dilution to the planet’s transit light curve due to the binary. The detector orientation was set to 178.7° in order to spatially separate the spectra. At this orientation, the spectra are separated by 7.9 pixels (compared to the FWHM of 1.1 pixel at $1.4 \mu\text{m}$). We describe the dilution correction in detail in § 3.3.

We also obtained photometric monitoring of the WASP-12 system to search for stellar activity. We acquired 428 out-of-transit R-band images over the 2011–

TABLE 1
SUMMARY OF PHOTOMETRIC OBSERVATIONS FOR WASP-12

Season	N_{obs}	Date Range (HJD - 2,400,000)	Sigma (mag)	Seasonal Mean (mag)
2011-2012	126	55824-56059	0.00325	-0.28048 ± 0.00029
2012-2013	169	56189-56439	0.00297	-0.28119 ± 0.00023
2013-2014	133	56463-56799	0.00316	-0.28253 ± 0.00027

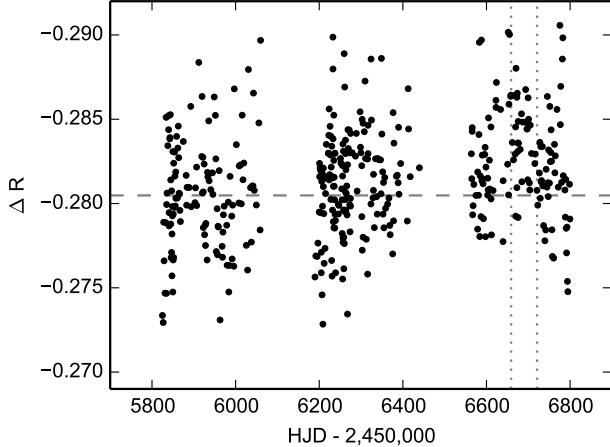


FIG. 2.— Nightly Cousins R -band photometric observations of WASP-12 over three observing seasons (points). The data are differential magnitudes relative to the average brightness of 17 comparison stars. The differential magnitudes have a mean of -0.281 and a standard deviation of 0.003. The horizontal dashed line indicates the mean brightness from the first observing season. The vertical dotted lines span the time range of the HST observations.

2012, 2012-2013, and 2013-2014 observing seasons with Tennessee State University’s Celestron 14-inch automated imaging telescope. We tested for variability from star spots by calculating differential magnitudes for WASP-12 relative to 17 comparison stars (shown in Figure 2). The standard deviation of the differential magnitudes is 0.3%, which is comparable to the photon noise for the data. The brightness of the star increases by approximately 0.001 mag per year. There are no significant periodicities between 1 and 200 days. To calculate the impact of star spots on our transmission spectrum, we took 0.3% as an upper limit for the variability of WASP-12. Variability of this amplitude could be produced by star spots 300 K cooler than the star’s effective temperature (6300 K; Hebb et al. 2009), covering 3% of the photosphere. Based on the formalism outlined in Berta et al. (2011) and Désert et al. (2011), we calculated that the maximum variation in transit depth due to spots with these properties is of order 10^{-5} , which is below the precision of our transit depth measurements.

2.2. Data Reduction

We reduced the *HST*/WFC3 data using the custom pipeline described in Kreidberg et al. (2014b). Example extracted spectra are shown in Figure 3. We bin the spectra into 24- and 15-pixel wide channels (for G102 and G141, respectively) to obtain a total of 13 spectrophotometric light curves at resolution $R \equiv \lambda/\Delta\lambda = 15 - 25$. We also sum the spectra over the full wavelength range to create broadband (“white”) light curves for each grism. The uncertainty in the flux per exposure is determined

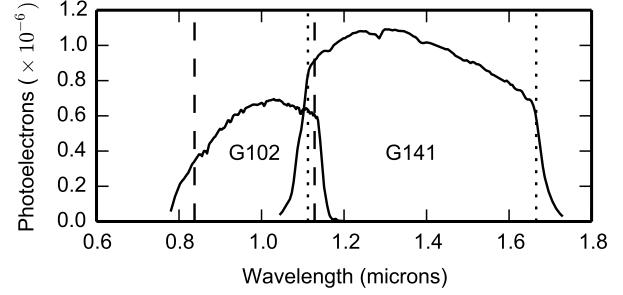


FIG. 3.— Example extracted spectra (solid lines). The wavelength ranges covered by the transmission spectra are indicated with dashed and dotted lines (for the G102 and G141 data, respectively). The uncertainties on the spectra are smaller than the plot linewidth.

by adding in quadrature the photon noise, the read noise (22 electrons per differenced image), and the error in the estimate of the background, which we determine from the median absolute deviation of the flux values for the background pixels.

We also extract spectra from the staring mode data to determine the wavelength-dependent flux ratio of WASP-12 BC to WASP-12 A. The spectra are spatially separated on the detector, so to extract them we fit a double Gaussian model to each column of the final detector read. We obtain final estimates and uncertainties for the spectra of WASP-12 A and WASP-12 BC by taking the mean and standard deviation of the spectra from all 10 exposures. These extracted spectra are each assigned a wavelength solution using the centroids of the stars in the direct image.

3. LIGHT CURVE FITS

3.1. Broadband Light Curves

We fit the broadband light curves with a transit model (Mandel & Agol 2002) and an analytic function to correct for instrument systematics. Transit light curve observations with *HST*/WFC3 have several well-documented systematic trends in flux with time, including visit-long slopes and orbit-long exponential ramps (Berta et al. 2012; Swain et al. 2013; Kreidberg et al. 2014b; Stevenson et al. 2014a). The first orbit of a visit has a larger amplitude ramp than subsequent orbits, so we follow established practice and do not use this initial orbit in the light curve fits. We also discard the first exposure from the remaining orbits, which improves the fit quality. We fit the remaining data (216 exposures for each grism) with a systematics model based on the `model-ramp` parameterization of Berta et al. (2012). The model $M(t)$ has the form:

$$M(t) = T(t) \times (c S(t) + v t_v) \times (1 - \exp(-a t_{\text{orb}} - b - D(t))). \quad (1)$$

In our fits, $T(t)$ is the transit model, or relative stellar flux as a function of time t (in BJD_{TDB}). The free parameters for the transit model are the planet-to-star radius ratio k , a linear limb darkening parameter u , the ratio of semi-major axis to stellar radius a/R_s , the orbital inclination i , and the time of mid-transit T_0 . The light curves have poor coverage of the planet's ingress, so we put priors on a/R_s and i to enable the measurement of T_0 . We use Gaussian priors with mean and standard deviation 2.91 ± 0.02 and $80.56 \pm 0.03^\circ$ (for a/R_s and i , respectively), based on estimates of those parameters from Stevenson et al. (2014c). We use an orbital period $P = 1.091424$ days and assume a circular orbit (Campo et al. 2011). The data from the two grisms are fit separately, but for each grism the three transits are fit simultaneously. We fit unique values of k and T_0 to each transit, but tie the values for u , a/R_s , and i over all the transits.

We fit the instrument systematics with a constant normalization term c , a scaling factor $S(t)$, a visit-long linear slope v , and an orbit-long exponential ramp with rate constant a , amplitude b , and delay $D(t)$. The timescale t_v corresponds to time relative to the expected transit midpoint for each visit, and t_{orb} is time since the first exposure in an orbit (both in BJD_{TDB}). The scaling factor $S(t)$ is equal to 1 for exposures with forward spatial scanning and s for reverse scanning; this accounts for a small offset in normalization between the scan directions caused by the upstream-downstream effect of the detector readout (McCullough & MacKenty 2012). The function $D(t)$ is equal to d for times t during the first fitted orbit and 0 elsewhere. A negative value for d implies that the ramp amplitude is larger in the first orbit than in subsequent orbits. The parameters u , a , b , and d were constrained to the same value for all the transits, whereas k , s , c , and v were allowed to vary between transits.

There are a total of 21 free parameters in the fits to each grism's broadband light curve. We estimated the parameters and their uncertainties with a Markov chain Monte Carlo (MCMC) fit to the data, using the `emcee` package for Python (Foreman-Mackey et al. 2013). For the best-fit light curves, we calculated the Durbin-Watson statistic to test for time-correlated noise. The values were 1.77 and 2.17 for the G102 and G141 white light curves, indicating that the residuals are uncorrelated at the 1% significance level. The reduced chi-squared values (χ^2_ν) for the best-fit light curves are 1.45 and 1.34, and the residuals are 164 and 115 ppm (37 and 25% larger than the predicted photon+read noise) for the G102 and G141 data, respectively.

To obtain conservative estimates of the errors for the white light curve transit parameters, we redid the MCMC analyses with the per-exposure uncertainties scaled by a constant factor (1.20 and 1.16) chosen to yield $\chi^2_\nu = 1$. We report the transit times from the white light curve fits in Table 2. These values extend the baseline of precise transit times by two years from the Stevenson et al. (2014c) measurements and will aid in testing the

TABLE 2
TRANSIT TIMES

Observation Start (UT)	T_c^a (BJD _{TDB})	Uncertainty (1σ)
Jan 01 2014	6659.07598	3.4E-4
Jan 16 2014	6674.35560	2.8E-4
Feb 05 2014	6694.00161	2.9E-4
Feb 15 2014	6703.82417	2.9E-4
Mar 02 2014	6719.10428	3.4E-4
Mar 04 2014	6721.28692	3.4E-4

^a We report the time of central transit T_c in BJD_{TDB} - 2,450,000.

possibility of perihelion precession for this system (first studied by Campo et al. 2011). The white light curves are consistent with the priors on a/R_s and i but do not yield improved values for those parameters due to the poor phase coverage of the transit ingress.

3.2. Spectroscopic Light Curves

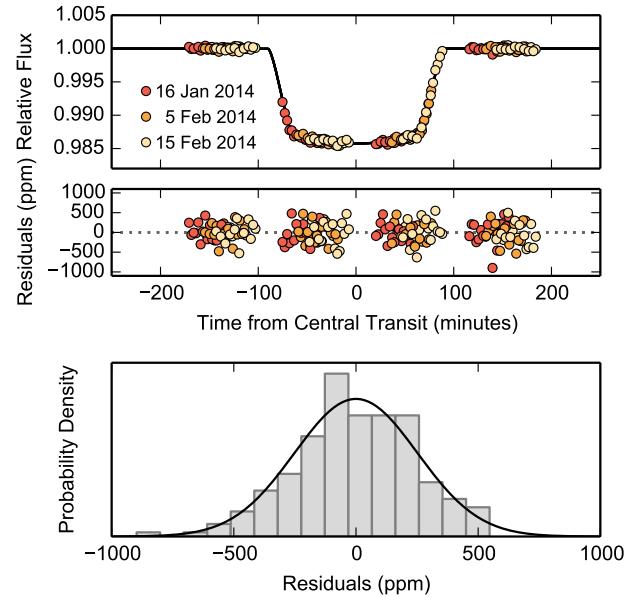


FIG. 4.— An example light curve fit to the 1.320 – 1.389 μm spectroscopic channel from the G141 grism. The top panel shows the best fitting model light curve (black line), overlaid with the systematics-corrected data (points). Residuals from the light curve fit are shown in the middle panel. The bottom panel shows a normalized histogram of the residuals compared to a Gaussian probability density with mean zero and standard deviation equal to the predicted photon+read noise (252 ppm).

We fit the spectroscopic light curves with a similar model as we use for the broadband data. The only differences are that we hold a/R_s and i fixed to the prior mean values and fix the mid-transit times to the white light curve best fit values. We also solve for a constant rescaling parameter for the photometric uncertainties in each spectroscopic channel to ensure that the reduced χ^2 for the light curve fits is unity. We chose to fix a/R_s , i , and T_c because they impact the mean spectroscopic transit depth only, not the relative depths. Changes in the mean transit depth do not significantly affect our

retrieval results, as the planet-to-star radius ratio is included as a free parameter in the atmospheric retrieval.

We fit each of the spectroscopic channel light curves independently. We achieve nearly photon-limited precision: the median rescaling factor for the photometric uncertainties is 1.1. The light curves do not exhibit statistically significant time-correlated noise, based on a Durbin-Watson test at the 1% significance level. An example spectroscopic light curve fit for the $1.320 - 1.389 \mu\text{m}$ channel from the G141 grism is shown in Figure 4. We show a pairs plot of the fit parameters for the same channel in Figure 5.

We tested an analytic model for the systematics that included a quadratic term for the visit-long trend. Previous analyses of WFC3 data have suggested that a quadratic model produces better light curve fits (Stevenson et al. 2014a). However, for this data set we find that the quadratic model is disfavored according to the Bayesian Information Criterion (BIC), with typical ΔBIC values of 10 compared to the linear model. In any case, the main consequence of adding a quadratic term is to shift the transmission spectrum up or down. This affects our estimate of the planetary radius, but the effect is small relative to the error introduced by uncertainties in the stellar radius.

We also explored modeling the instrument systematics with the **divide-white** technique, which has been applied successfully to several other data sets (Stevenson et al. 2014c; Kreidberg et al. 2014b; Knutson et al. 2014). This method assumes the systematics are independent of wavelength. For the WASP-12 data, however, this assumption is not appropriate. WFC3 instrument systematics are known to depend on detector illumination (Berta et al. 2012; Swain et al. 2013), and in our data the mean pixel fluence varies by 30% between the spectroscopic channels (see Figure 3).

To illustrate the dependence of the systematics on the illumination level, we show in Figure 6 the systematics decorrelation parameters from the analytic model as a function of light curve normalization c . The value c represents the baseline flux level in each spectroscopic light curve. This value is directly proportional to the mean detector illumination in the channel. We note several qualitative trends in the decorrelation parameters: with increasing illumination, the visit-long slope decreases, the delay term increases toward zero, and the ramp rate increases.

The upshot of these trends is that there are residual systematics for the **divide-white** light curve fits that are correlated with detector illumination. We therefore only report the results for the analytic model.

3.2.1. Limb Darkening Models

We tested fixing the limb darkening to values predicted by stellar models. We used theoretical quadratic limb darkening coefficients from both PHOENIX and Kurucz models (Hauschildt et al. 1999; Castelli & Kurucz 2004) generated for Stevenson et al. (2014c). These theoretical coefficients yielded lower quality light curve fits than we obtained from empirically estimating a linear limb darkening parameter. For both PHOENIX and Kurucz model coefficients, the light curve residuals exhibited systematic trends near ingress and egress, and the typical reduced chi-squared values increased. The poor match of theoretical limb darkening coefficients to our data may arise from inaccurate assumptions about the stellar composition, or inaccuracies in the models themselves. Similar disagreement with model limb darkening has been seen for other high quality data (e.g. Knutson et al. 2007).

Since incorrect limb darkening coefficients can introduce systematic bias in the measured transit depths, we chose to estimate the limb darkening coefficients directly from the data. Our light curves are not precise enough to constrain a two-parameter limb darkening model, so we instead fit for a single linear parameter in each channel. The data are sufficiently precise to distinguish between the transit depth and the limb darkening coefficient, as evidenced by the lack of correlation between those parameters in the pairs plot shown in Figure 5.

3.3. Correction of Dilution from Stellar Companions

The transit light curves are affected by dilution from WASP-12 A's two companion stars, WASP-12 BC, and from the planet's nightside emission. Following Stevenson et al. (2014c), we calculated a corrected transit depth δ' for each spectroscopic channel using the formula

$$\delta' = \delta(1 + \alpha_{\text{Comp}} + \alpha_p) \quad (2)$$

where δ is the measured transit depth, α_{Comp} is the ratio of flux from WASP-12 BC to WASP-12 A, and α_p is the ratio of the planet's nightside emission to the flux of WASP-12 A. We calculated α_p using the same model as Stevenson et al. (2014c). We determined α_{Comp} empirically using the spectra extracted from the staring mode observations. Each visit is assigned a unique α_{Comp} value to account for differences in the orientation of the spectra on the detector. We plot the dilution factor as a function of wavelength in Figure 7 (calculated by interpolating the companion spectrum to the same wavelength scale as the WASP-12 A spectrum). The measured dilution is roughly 10% larger than the values used by Stevenson et al. (2014c); this difference could result from systematic uncertainty in the aperture photometry used to scale the dilution (Stevenson et al. 2014c).

To obtain final transit depths, we take the weighted average of the corrected transit depths from each visit. Uncertainties in the dilution factor corrections are propagated through to the final transit depth uncertainties. We report the corrected transit depth measurements in Table 3, and we show the transmission spectrum in Figure 8. The corrected transit depths are consistent between the visits, indicating that the estimated uncertainties are appropriate, and confirming that stellar activity does not significantly influence the measured spectrum.

4. COMPARISON WITH PREVIOUS RESULTS

There are several other high-precision transit depth measurements for WASP-12b in addition to the WFC3 spectrum we report here. The most precise of these are spectroscopy between 0.34 and $0.94 \mu\text{m}$ from *HST/STIS* (Sing et al. 2013), spectroscopy between 0.73 and $1.00 \mu\text{m}$ from *Gemini/GMOS* (Stevenson et al. 2014c), a staring mode *HST/WFC3* G141 spectrum from Swain et al. (2013), and photometry at 3.6 and $4.5 \mu\text{m}$ from *Spitzer/IRAC* (Cowan et al. 2012). Figure 9 shows past optical/near-IR transmission spectrum measurements compared to our results.

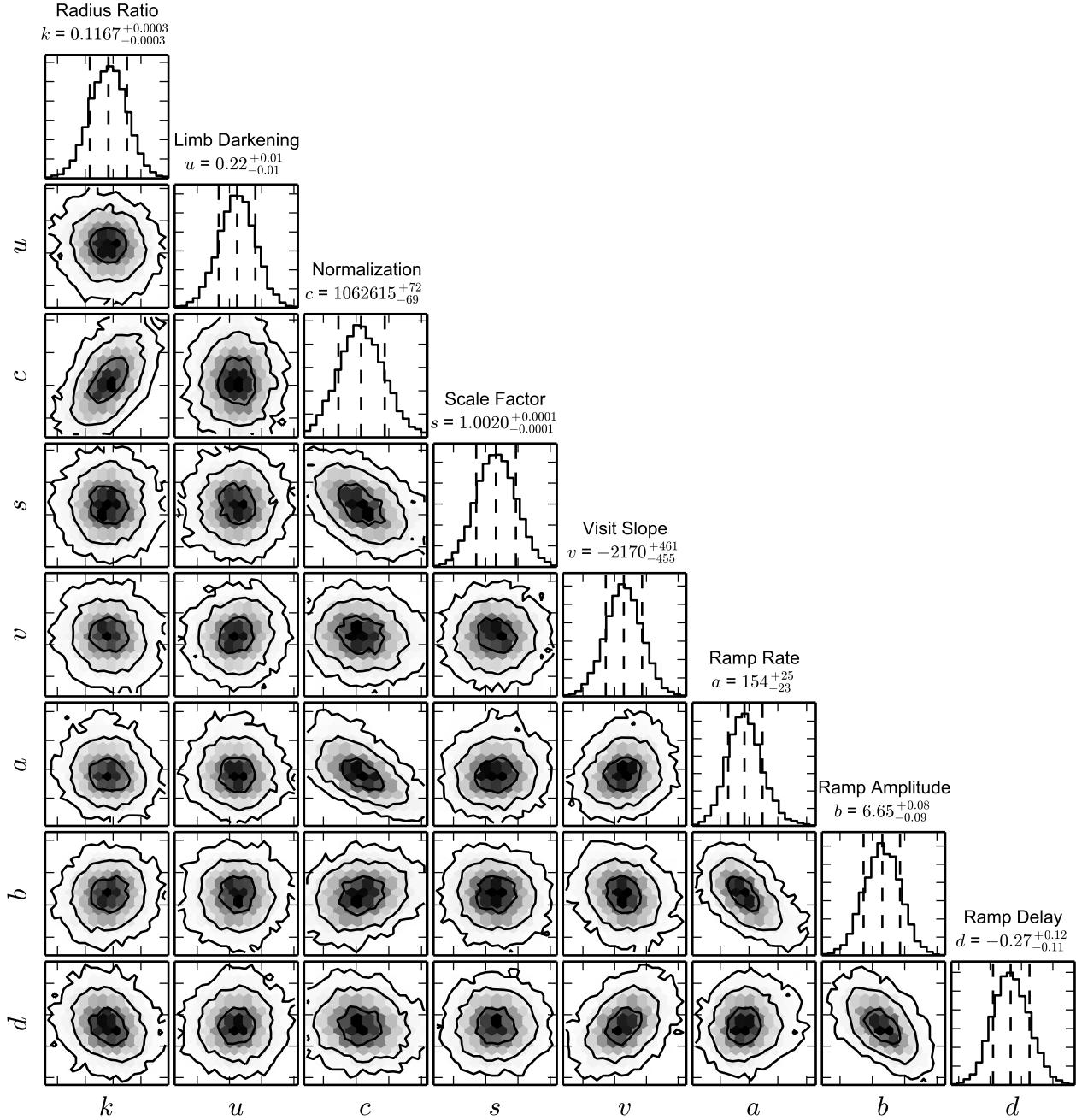


FIG. 5.— A pairs plot showing distributions of MCMC fit parameters for the $1.320 - 1.389 \mu\text{m}$ light curve. The off-diagonal panels show marginalized posterior probability for pairs of parameters, with 1, 2, and 3σ credible intervals indicated with black contours. The gray shading corresponds to probability density (darker for higher probability). The panels on the diagonal show distributions of each parameter marginalized over the other model parameters, with the median and 68% credible interval marked with dashed lines. The planet-to-star radius ratio, k , is not strongly correlated with any of the other fit parameters. For parameters that are allowed to vary between transit observations (k , c , s , and v), we show distributions for the 16 Jan 2014 transit.

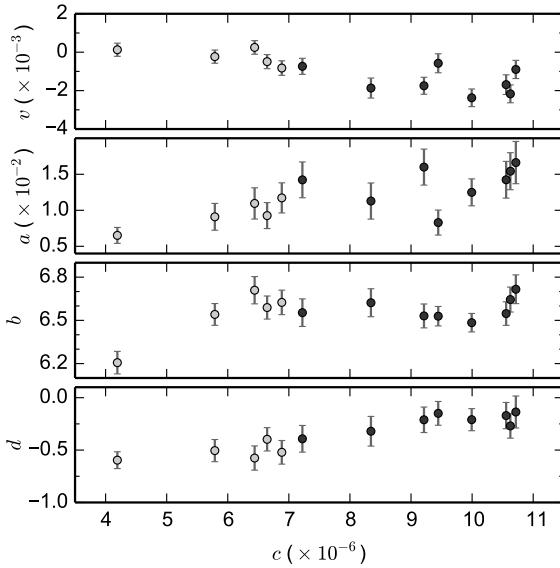


FIG. 6.— Systematics decorrelation parameters from the light curve fits plotted as a function of normalization c . We show (from top to bottom) the visit-long slope v , ramp rate constant a , ramp amplitude b , and ramp delay d . The normalization constant c is linearly proportional to the per-pixel flux in a spectroscopic channel. Measurements from the G102 and G141 data are shown in light and dark gray, respectively. The error bars indicate 1σ uncertainties from an MCMC fit to the light curves using the analytic model for instrument systematics.

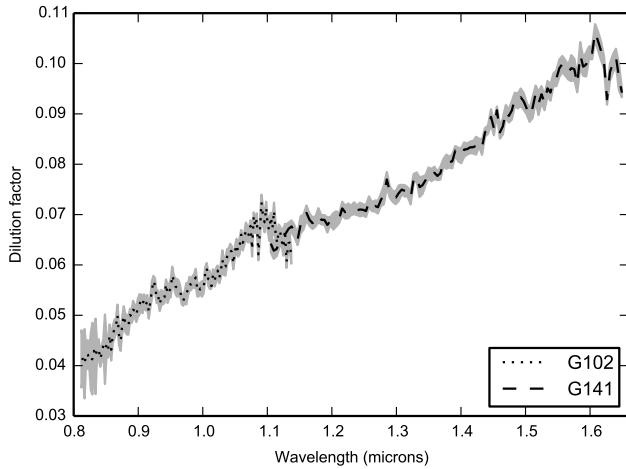


FIG. 7.— The dilution factor α_{Comp} , determined from the ratio of flux from WASP-12 BC to WASP-12 A. The median dilution (over the 10 staring mode observations) is shown with black lines (dotted for G102 and dashed for G141). The gray shaded region indicates 1σ uncertainty, determined from the median absolute deviation. To calculate the dilution and uncertainty, we interpolate the spectra from each grism onto a common wavelength scale.

Our transit depth measurements are consistent with the staring mode spectrum from *HST*/WFC3 (Swain et al. 2013; Mandell et al. 2013; Stevenson et al. 2014c), modulo a constant offset in the absolute transit depths. We chose not to incorporate this data set in our analysis, however. The staring mode data has a low duty cycle and increases the amount of in-transit exposure time by just 10% over the total from our three spatial scan ob-

TABLE 3
TRANSIT FIT PARAMETERS

Bandpass ^a (μm)	$(R_p/R_\star)^2$ ^b (%)	Limb-darkening u	rms (ppm)
0.838 – 0.896	1.4441 ± 0.0069	0.27 ± 0.02	349
0.896 – 0.954	1.4422 ± 0.0055	0.28 ± 0.01	300
0.954 – 1.012	1.4402 ± 0.0052	0.27 ± 0.01	296
1.012 – 1.070	1.4428 ± 0.0051	0.26 ± 0.01	301
1.070 – 1.129	1.4391 ± 0.0053	0.25 ± 0.01	299
1.112 – 1.182	1.4386 ± 0.0047	0.26 ± 0.01	267
1.182 – 1.251	1.4365 ± 0.0045	0.26 ± 0.01	267
1.251 – 1.320	1.4327 ± 0.0041	0.21 ± 0.01	253
1.320 – 1.389	1.4582 ± 0.0040	0.22 ± 0.01	242
1.389 – 1.458	1.4600 ± 0.0043	0.18 ± 0.01	255
1.458 – 1.527	1.4530 ± 0.0045	0.20 ± 0.01	270
1.527 – 1.597	1.4475 ± 0.0058	0.16 ± 0.02	332
1.597 – 1.666	1.4332 ± 0.0055	0.16 ± 0.02	321

^a The measurements between 0.838 – 1.129 μm are from the G102 grism; those from 1.112 – 1.666 μm are from the G141 grism.

^b Transit depths corrected for dilution from companion stars and planet nightside emission. The light curve fits had a/R_s and i fixed to 2.91 and 80.56° , respectively.

servations. Moreover, there is uncertainty about the best instrument systematics model for this data set (Stevenson et al. 2014c). Since the marginal improvement in measurement precision is counteracted by increased systematic uncertainty, we chose to focus on the transmission spectrum derived from the spatial scan data only.

Our spectrum agrees less well with the ground-based transmission spectrum from *Gemini*/GMOS (Stevenson et al. 2014c). To compare the WFC3 and GMOS results, we computed χ^2 values for the data using the best fit model spectrum from the FREE retrieval (described in § 5.1) over the wavelength range where the spectra overlap (0.86 – 1.00 μm). We allowed a free offset in the model for the GMOS data to account for differences in the absolute measured transit depths. We find that the GMOS data are inconsistent with the model at $> 4\sigma$ ($\chi^2 = 38.7$ for 9 degrees of freedom). Based on the track record of precise, reproducible transmission spectra from WFC3 that are well-fit by theoretical models (e.g. Kreidberg et al. 2014b,a; Stevenson et al. 2014d), we trust the reliability of the WFC3 measurements over those from GMOS. Ground-based transit observations exhibit strong time-correlated noise that makes estimating accurate confidence intervals challenging. Additional GMOS transit spectroscopy observations would help identify the number of repeated observations with this instrument needed to make robust measurements.

The *Spitzer* transit depths measurements are approximately 0.1% smaller than the mean WFC3 transit depth. Based on the random errors alone, this difference is significant. However, the *Spitzer* data have significant systematic uncertainties, because the relative transit depths measured with IRAC channels 1 and 2 can shift by $> 0.1\%$ depending on the aperture size used in the data reduction. There is no obvious optimal choice of aperture (see Figure 17, Stevenson et al. 2014c). We therefore chose not to incorporate the *Spitzer* results in our analysis.

5. RETRIEVAL OF THE ATMOSPHERIC PROPERTIES

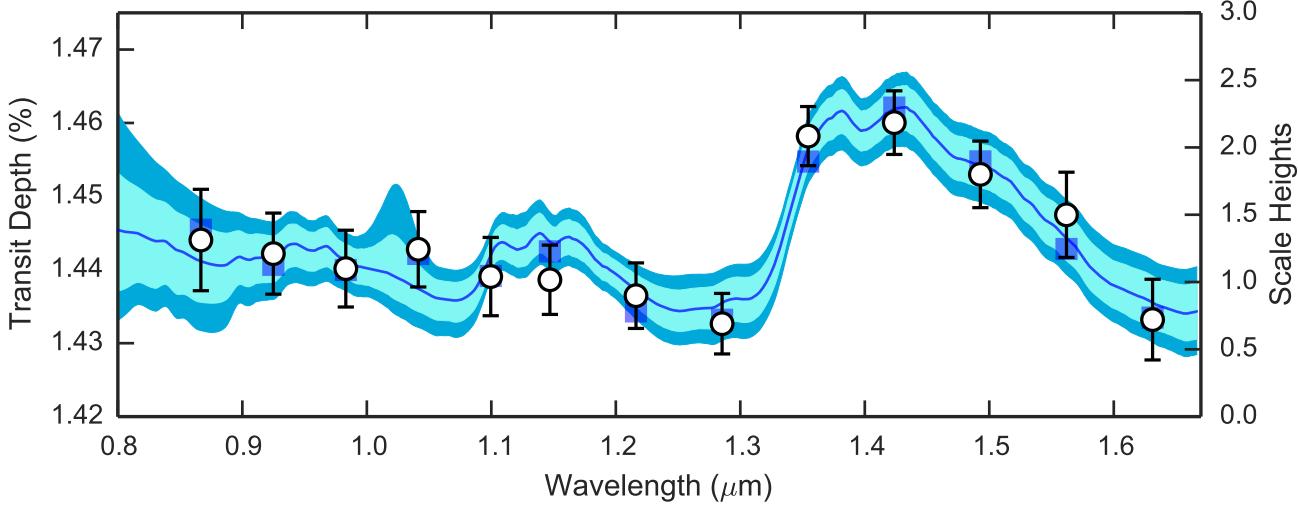


FIG. 8.— The transmission spectrum of WASP-12b measured with *HST*/WFC3 (points). The error bars on the transit depths are 1σ uncertainties from an MCMC fit to the light curves. We show the best fit model binned to the resolution of the data (blue squares). The shaded regions indicate 1- and 2σ credible intervals in the retrieved spectrum (medium and light blue, respectively), relative to the median fit (dark blue line). The secondary y-axis labels indicate the atmospheric scale height. One scale height corresponds to 470 km (for a temperature of 1400 K). The increase in transit depth near $1.4\mu\text{m}$ corresponds to an H_2O bandhead.

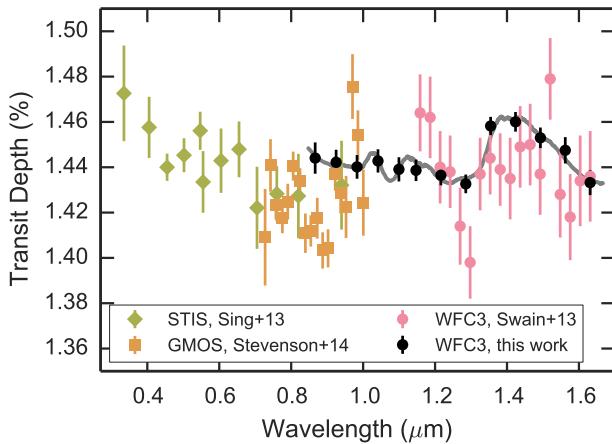


FIG. 9.— High-precision optical/near-IR transit depth measurements for WASP-12b. We show measurements from *HST*/STIS (green diamonds), *Gemini*/GMOS (orange squares), and *HST*/WFC3 (from this work, black circles; from staring mode data, pink circles). The best fit model from the CHIMERA FULL retrieval is indicated by the gray line. Note that this model was fit to the WFC3 data only. There may be small offsets between data sets due to different corrections for the companion star dilution and the challenge of measuring absolute transit depths in the presence of systematic errors (Stevenson et al. 2014c,a).

Given the high precision of the WFC3 transmission spectrum, we can put more powerful constraints on the planet's atmosphere than has been possible with past measurements. No previous data set for WASP-12b has shown conclusive spectroscopic evidence for molecular absorption, but the WFC3 spectrum has a noticeable increase in transit depth near the center of the water absorption band at $1.4\mu\text{m}$ (see Figure 8).

To quantify the water abundance and other atmospheric properties based on the transmission spectrum, we performed a retrieval using the CHIMERA suite (Line et al. 2013b, 2014). Modifications to the retrieval code

and description of the transmission forward model are provided in Line et al. (2013a); Stevenson et al. (2014d); Kreidberg et al. (2014a); Swain et al. (2014); Diamond-Lowe et al. (2014). We tested two different model parameterizations with CHIMERA. The first is the widely adopted approach of retrieving molecular abundances without any constraints from chemistry (e.g. Madhusudhan & Seager 2009). For our second approach, we developed a new parameterization that retrieves C/O and atmospheric metallicity under the assumption of chemical equilibrium. We refer to these methods as the FREE approach and the CHEMICALLY-CONSISTENT (C-C) approach, respectively. Both methods are described in more detail below. We also present a comparison with results from the NEMESIS retrieval code (Irwin et al. 2008).

We elected to focus on the WFC3 data for our retrievals, based on the caveats for other data described in § 4. We treated the spectra from the two WFC3 grisms as a single data set, with no offset in transit depth between them. The transit depths are consistent between the grisms because the influence of stellar activity is below the measurement precision (see § 2.1), so no offset is needed. We also tested including STIS data in some of the retrievals (see § 5.1.2), but found that it does not affect the main conclusions of this work.

5.1. FREE Retrieval with CHIMERA

The FREE parameterization retrieves molecular abundances, cloud and haze properties, and the altitude-independent scale height temperature.

Our model explored the dominant molecular opacities expected for a solar composition gas at the temperatures and pressures probed by the observations. These are H_2O , TiO , Na , and K over the wavelength range of the WFC3 spectrum (Burrows & Sharp 1999; Fortney et al. 2008). The remaining gas is assumed to be a solar composition mixture of H_2 and He . We discuss results for

including additional molecular species in § 5.1.2.

In addition to the molecular opacities, we included opacity from clouds and haze. Previous analyses of WASP-12b's transmission spectrum have suggested that hazes are present in the atmosphere, but the haze composition is poorly constrained (Sing et al. 2013; Stevenson et al. 2014c). We therefore used a flexible model that includes a power law haze and an opaque gray cloud deck. We modeled clouds as a gray opacity source that masks transmission through the atmosphere below a fixed pressure level P_c . The haze opacity was parameterized by $\sigma(\lambda) = \sigma_0(\lambda/\lambda_0)^\gamma$, where the scattering amplitude σ_0 and slope γ are free parameters (as in Lecavelier Des Etangs et al. 2008). The scattering is presumed to happen throughout the entire atmosphere as if it were a well mixed gas. The scattering amplitude is a scaling to the H₂ Rayleigh scattering cross section times the solar H₂ abundance at 0.43 μm . More sophisticated models for clouds and haze would require additional free parameters that are not justified by the precision of our data. Our simple parameterization is sufficient to capture degeneracies between clouds/haze and the water abundance, which is the primary goal of this investigation.

We also retrieved a “scale height temperature” parameter T_s . We assumed the atmosphere is isothermal, with temperature equal to T_s at all pressures. We tested fitting for a more complex temperature-pressure profile but found it did not affect our results because the transmission spectrum is only weakly sensitive to the atmosphere's thermal structure (see Barstow et al. 2013, 2014).

The final free parameter in the retrieval was a scale factor for the planet's 10 bar radius. We assumed a baseline planet radius of 1.79 R_J and a stellar radius of 1.57 R_\odot from Hebb et al. (2009). The planet radius was scaled by a factor R_{scale} to account for uncertainty in the pressure level in the atmosphere at a given radius. To first order, the effect of this scaling is to shift the model transit depths by a constant factor. A second order effect is that scaling the radius changes the amplitude of spectral features (see Equation 1 in Lecavelier Des Etangs et al. 2008).

In sum, the retrieval had 8 free parameters: the abundances of H₂O, Na+K (fixed at the solar abundance ratio), and TiO, as well as clouds and haze, the scale height temperature, and the planet radius scale factor. We refer to this 8-parameter fit as the FULL model. The best fits for the FULL model and nested models within it are shown in Figure 10. The best-fit models are those that produced the lowest χ^2 values in the MCMC. We list the retrieved water abundances, temperatures, and best fit χ^2 values for these models in Table 4. The distributions of retrieved parameters from the FULL model are shown in Figure 11. We note that all of the nested models are nearly indistinguishable near the water absorption feature at 1.4 μm . The best constrained molecular abundance is that of water: we retrieved a water volume mixing ratio (VMR) of $1.6 \times 10^{-4} - 2.0 \times 10^{-2}$ at 1σ , which is consistent with expectations for a solar composition gas. The other molecular abundances are not as well constrained. We retrieved a 3σ upper limit on the VMR of TiO equal to 2×10^{-4} , and the abundance of Na+K is unbounded. The cloud and haze properties are also poorly constrained. We discuss the implications of

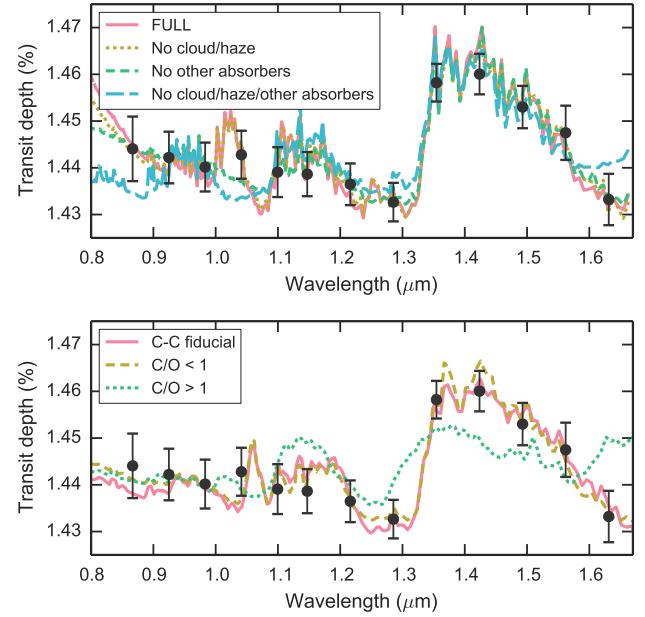


FIG. 10.— Best fits for different assumed models (lines) compared to the WFC3 transmission spectrum (points). Models are binned to a wavelength resolution of 0.01 μm . Results from the FREE and C-C parameterizations are shown in the top and bottom panels, respectively.

these measurements for the atmosphere composition in detail in § 6.

5.1.1. Detection Significances

We performed nested model selection to identify how strongly each opacity source is detected in the spectrum. The standard Bayesian approach for comparing models is to use Bayes factors, which can be converted to detection significances (Sellke et al. 2001; Trotta 2008). See also Benneke & Seager (2013) for an application of nested Bayesian model comparisons as applied to super-Earth atmospheres. Bayes factors are the ratio of the Bayesian evidence (marginal likelihood) of the two models under consideration. The evidence is a multidimensional integral over the entire posterior volume, a non-trivial calculation, so here we use two methods to approximate the integral. The first is the Numerical Lebesgue Algorithm described in Weinberg (2012) (see also Swain et al. 2014, for an application to exoplanet atmospheres and a comparison to the $\Delta\chi^2$ test). The second is the Laplace approximation (Kass & Raftery 1995; Cornish & Littenberg 2007). Most methods for computing Bayesian evidence diverge for low detection significances, but they agree well for highly significant detections (see Fig. 3 from Cornish & Littenberg 2007).

Using these techniques, we determined the detection significances for the following parameter combinations: H₂O, absorbers other than H₂O, clouds and haze combined, clouds only, haze only, and all opacity sources other than H₂O. We computed the detection significance for each nested model by calculating its Bayes factor relative to the FULL model. The results are shown in Table 4. If a parameter has a low detection significance value ($\lesssim 3\sigma$), that parameter does not provide a statistically significant improvement to the model fit.

We find that H₂O is detected at high confidence in the

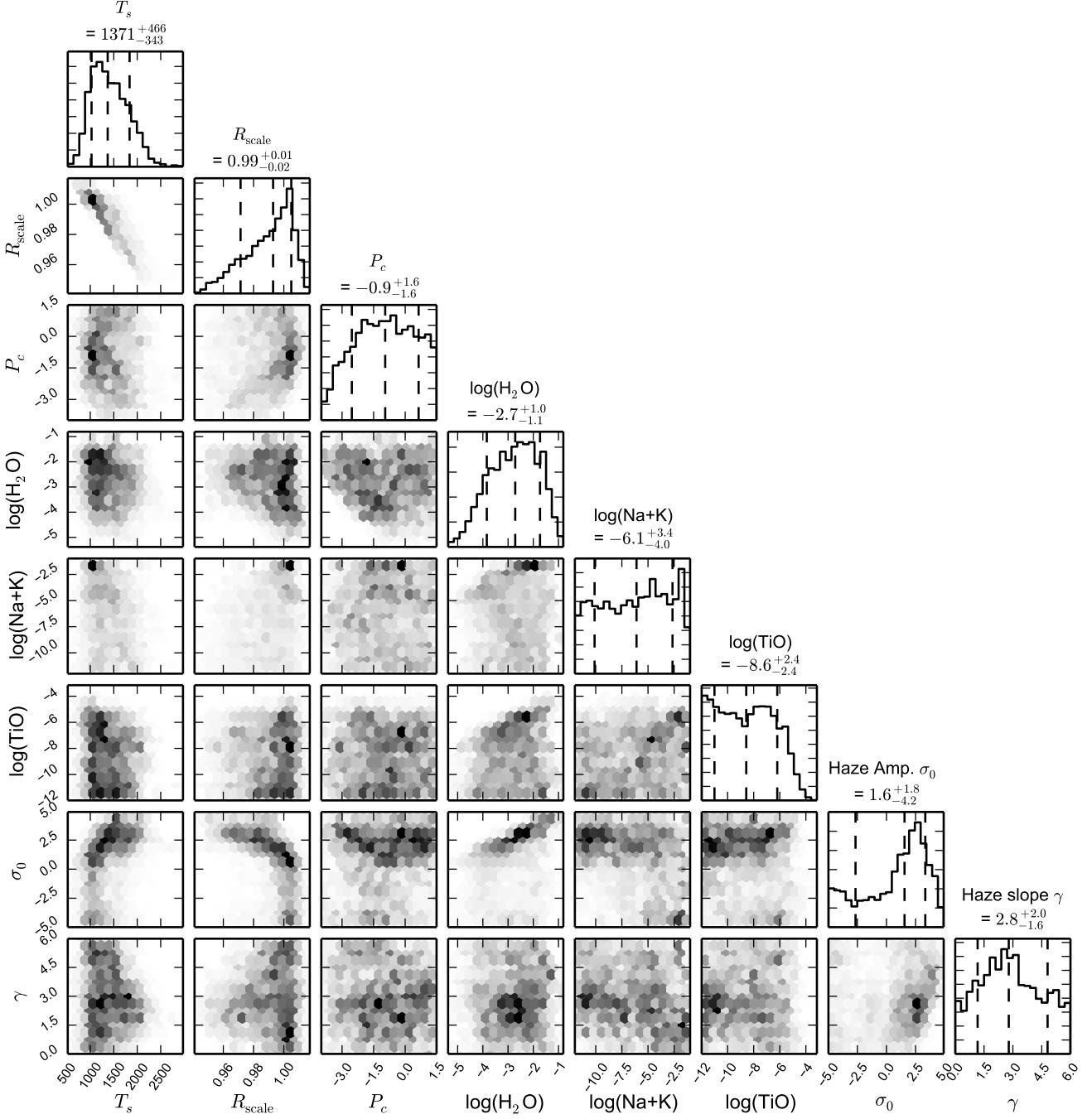


FIG. 11.— A pairs plot of the distribution of parameters retrieved with CHIMERA for the FULL model. The parameters are the scale height temperature T_s (in Kelvin), the planet radius scale factor R_{scale} , the cloud-top pressure P_c (in log bars), the logarithm (base 10) of the molecular abundances, and the haze scattering amplitude σ_0 and slope γ . The off-diagonal plots show marginalized posterior probability density for pairs of parameters (darker shading corresponds to higher probability). The diagonal plots show marginalized posterior distributions for individual parameters, with the median and 68% credible interval marked with dashed lines.

spectrum (7σ). We also find that the presence of all other opacity sources combined (clouds, haze, TiO, Na and K) is significant at $> 3\sigma$. However, no other combination of these opacities is significantly detected (e.g., the detection significance for clouds alone is $< 2\sigma$). The reason for this is that the parameters are degenerate. For example, both clouds and haze both have the effect of truncating the height of the water feature. Similarly, haze, the wings of the alkali metal lines, and presence of TiO can all contribute to the rise in transit depths

towards the blue end of the spectrum. Therefore, if a few of these opacity sources are removed, the others can compensate. However, removing all opacities besides water results in a significantly poorer model fit (as can be seen in Figure 10). We therefore focus on results from the FULL model because it accounts for the likely presence of some combination of these other species and the degeneracies they may have with H_2O .

5.1.2. Retrievals with Additional Molecules/Data

TABLE 4
FREE RETRIEVAL RESULTS

Scenario ^a	Water abundance ^b (VMR)	Temperature ^b (Kelvin)	χ^2 ^c	Weinberg significance ^d (σ)	Laplace significance (σ)
FULL ($T_s, R_{\text{scale}}, P_c, \text{H}_2\text{O}, \text{TiO}, \text{Na+K}, \sigma_0, \gamma$)	$1.5 \times 10^{-4} - 2.2 \times 10^{-2}$	1040 – 1870	2.82	–	–
No water ($T_s, R_{\text{scale}}, P_c, \text{TiO}, \text{Na+K}, \sigma_0, \gamma$)	–	130 – 1560	47.0	6.9	7.1
No cloud, haze, other absorbers ($T_s, R_{\text{scale}}, \text{H}_2\text{O}$)	$1.1 \times 10^{-5} - 1.5 \times 10^{-2}$	730 – 1170	11.1	3.2	4.6
No other absorbers ($T_s, R_{\text{scale}}, P_c, \text{H}_2\text{O}, \sigma_0, \gamma$)	$6.9 \times 10^{-5} - 2.9 \times 10^{-2}$	1090 – 1890	3.59	1.2	2.6
No haze ($T_s, R_{\text{scale}}, P_c, \text{H}_2\text{O}, \text{TiO}, \text{Na+K}$)	$5.8 \times 10^{-5} - 8.3 \times 10^{-3}$	910 – 1470	2.87	1.6	2.5
No cloud ($T_s, R_{\text{scale}}, \text{H}_2\text{O}, \text{TiO}, \text{Na+K}, \sigma_0, \gamma$)	$6.6 \times 10^{-5} - 2.0 \times 10^{-4}$	1000 – 1840	2.86	–	2.0
No cloud/haze ($T_s, R_{\text{scale}}, \text{H}_2\text{O}, \text{TiO}, \text{Na+K}$)	$2.4 \times 10^{-5} - 5.5 \times 10^{-3}$	860 – 1430	2.95	–	1.2

^a Indicates the parameters that are removed from the FULL model. The remaining parameters are shown in parentheses.

^b The range of values corresponds to the 68% credible interval centered on the median retrieved value.

^c χ^2 values are calculated for the best fit model.

^d Scenarios where the detection significance is undefined (because the Bayes factor is less than 1) are marked by –. The detection significance for the FULL model is also undefined because significances are defined relative to it.

We explored many different combinations of models and data sets before we arrived at the analysis presented above. One option we considered was to fit a WFC3 spectrum with smaller wavelength bins. We created a “high-resolution” spectrum with $R = 55 - 70$, in contrast to the “low-resolution” data ($R = 15 - 25$) we used for our final modeling. The spectra are qualitatively similar, but the models fit the low-resolution data better ($\chi^2_{\nu} = 0.6$ versus 1.5). The high-resolution spectrum has larger random scatter around the best fit model, but no systematic trends indicating that the model is missing a particular physical effect. The scatter could be due to undiagnosed systematics at the pixel level in the data, which is remedied by increasing the number of pixels that contribute to each spectral channel.

The retrieved water abundances for both resolutions are very similar ($1.5 \times 10^{-4} - 2.2 \times 10^{-2}$ and $1.2 \times 10^{-4} - 1.8 \times 10^{-2}$ for low- and high-resolution). On the other hand, the best fit scale height temperatures differ by 1.4σ (1370^{+470}_{-340} versus 2020^{+320}_{-340} for low/high). This difference is due to slight changes in the morphology of the water feature between the data sets. It illustrates how retrieval results can be biased or have underestimated uncertainties when the model does not provide a good fit to the data, either because of missing model physics or underestimated data error bars.

Another test we performed was to fit the combined transmission spectrum from *HST*/STIS and WFC3. The STIS spectrum covers optical wavelengths and exhibits a linear increase in transit depth from red to blue (see Figure 9). Incorporating the STIS data in the retrieval increased the detection significance of clouds and haze to $> 3\sigma$ (because of the rise in transit depth towards the blue), but otherwise did not significantly change the results.

We also tested including additional absorbing species in the retrieval. Previous work has suggested WASP-12b has a carbon-rich atmosphere composition (e.g. Madhusudhan et al. 2011b), so we wanted to confirm that our estimate of the water abundance is not biased by only including absorbers expected for solar composition. We therefore ran a retrieval including all the major opacity sources expected for either oxygen-rich or carbon-rich compositions, including H_2O , CO , CO_2 , NH_3 , TiO , VO , Na , K , CH_4 , C_2H_2 , HCN , H_2S , FeH , N_2 , and collisionally-induced H_2/He absorption (Burrows &

Sharp 1999; Fortney et al. 2008; Madhusudhan 2012). For this retrieval, we also fit the STIS data and used the higher-resolution WFC3 spectrum. The retrieved water abundance of $2.5 \times 10^{-4} - 2.0 \times 10^{-2}$ is nearly identical to the results we obtained from the FULL model. No individual absorbers were significantly detected besides H_2O .

Our conclusion from all the scenarios we fit is that the constraints on the water abundance are not significantly impacted by the choice of data sets or the absorbers included in the modeling.

5.2. CHEMICALLY-CONSISTENT Retrieval with CHIMERA

In addition to retrievals with the FREE parameterization (described in § 5.1), we also developed a reparameterized model to retrieve atmospheric properties that are consistent with chemical equilibrium. Rather than varying the mixing ratios of individual gases, we made C/O and metallicity free parameters and computed thermochemical equilibrium abundances along the temperature/pressure (T/P) profile for major species. We retrieved the T/P profile rather than a scale height temperature in order to explore any additional degeneracies due to the profile shape. The T/P profile is modeled with a 5-parameter analytic function (Line et al. 2013b) which produces physically realistic thermal profiles consistent with radiative equilibrium (e.g. Guillot 2010; Robinson & Catling 2012; Heng et al. 2012; Parmentier & Guillot 2014). We calculated chemical profiles for H_2 , He , H_2O , CH_4 , CO , CO_2 , NH_3 , H_2S , PH_3 , C_2H_2 , HCN , Na , K , FeH , TiO , VO , and N_2 to cover the full range of gases that could contribute significant opacity in the near-IR. The equilibrium abundances are computed using the NASA Chemical Equilibrium with Applications code (McBride & Gordon 1996; Moses et al. 2011; Line et al. 2011). The chemical profiles were fed into the radiative transfer model to calculate model transmission spectra for comparison with our data. The model also included opacity from clouds and haze (using the same formalism described in § 5.1), and the reference radius R_p , for a total of 11 free parameters. While we reduced the number of free absorber parameters to just C/O and metallicity, the overall number of free parameters increased due to the more flexible T/P profile.

We ran retrievals using three different priors for C/O. The fiducial case had an uninformative prior constraint

TABLE 5
C-C RETRIEVAL RESULTS

Scenario	C/O	Metallicity (\times solar)	Water abundance (1σ)	Temperature ^a (K) (1σ)	χ^2
Fiducial	0.2 – 0.7	0.3 – 20	9.3×10^{-5} – 5.3×10^{-3}	1090 – 1760	2.43
O-rich	0.1 – 0.7	0.3 – 30	1.2×10^{-4} – 1.5×10^{-2}	1090 – 1890	2.75
C-rich	2.9 – 51	0.2 – 80	1.5×10^{-5} – 8.6×10^{-3}	120 – 550	29.2

^a The temperature range corresponds to a pressure level of 1 mbar.

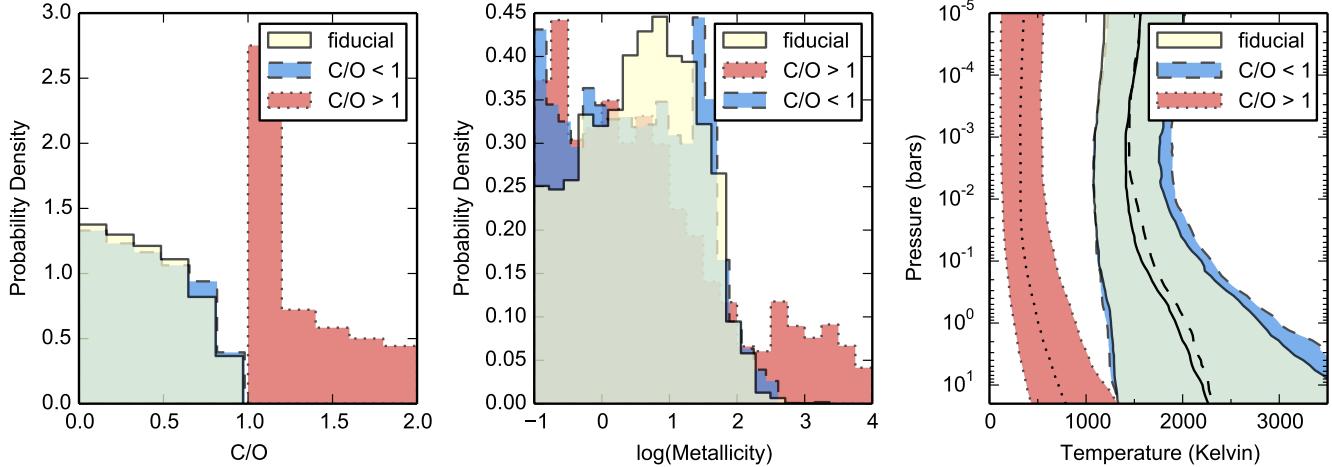


FIG. 12.— Retrieval results from the C-C parameterization. The left and middle panels show the marginalized distribution for C/O and metallicity. The right panel shows the median (black lines) and 1σ range (shaded regions) of 1000 randomly sampled temperature-pressure profiles for each scenario. Yellow, blue, and red shading correspond to constraints from the fiducial scenario (an uninformative prior on C/O), the oxygen-rich scenario ($C/O < 1$), and a carbon-rich scenario ($C/O > 1$). The distribution of C/O values for the carbon-rich model is normalized to have a probability mass of unity over the plotted range; however, the distribution has an extended tail toward higher C/O values that is not shown.

on C/O. We also ran retrievals that constrained the atmospheric composition to be either oxygen-rich or carbon-rich. For the O-rich scenario, the prior probability was set to zero for C/O values greater than unity. Correspondingly, the C-rich scenario had zero prior probability for $C/O < 1$. We show the best fit models for all three cases in Figure 10. Figure 12 shows marginalized distributions of C/O and metallicity for each scenario, as well as constraints on the temperature-pressure profiles. Table 5 gives the χ^2 values of the best fits in each scenario. It also lists the 68% credible intervals for C/O, metallicity, and the temperature and water abundance at 1 mbar pressure.

The fiducial and oxygen-rich scenarios give nearly identical constraints on C/O and the atmospheric metallicity. Even though the fiducial model has an uninformative prior on C/O, 100% of the retrieved C/O values are less than unity for this case (suggesting the result is data-driven rather than prior-driven). Both scenarios yield constraints on the molecular abundances and thermal profile that agree well with results from the FULL model. The median retrieved water abundances are 6.3×10^{-4} and 1.0×10^{-3} for the fiducial and O-rich scenarios, respectively. The median 1 mbar temperatures are 1410 and 1450 K. We retrieved metallicities in the range $0.3 – 30 \times$ solar (at 1σ). The retrieved cloud and haze properties for these scenarios are consistent with the FULL model at 1σ .

By contrast, the carbon-rich scenario produced signif-

icantly different atmospheric properties. To reproduce the water absorption feature in the spectrum, the temperature was driven to much lower values (the median is 320 K at 1 mbar). These low temperatures allow for higher water abundances (see Figure 13), though they are unlikely for the terminator region given how highly irradiated the planet is. The median water abundance, 1.4×10^{-4} , is comparable to that for the O-rich and fiducial scenarios. On the other hand, the median methane abundance increases to 2.4×10^{-3} , versus 1.3×10^{-11} for the fiducial case. These differences have two main effects on the model transmission spectra. One is that absorption features have smaller amplitude because lower temperatures decrease the atmospheric scale height. The second effect is that methane absorption is present in the spectrum. It is especially noticeable in the window between water features at $> 1.6 \mu\text{m}$ wavelengths (see the best-fit C-rich model in Figure 10). The consequence of these changes is that C-rich models do not fit the measured spectrum as well as O-rich models. We discuss the strength of the evidence for one scenario over the other in quantitative detail in § 6.

As a test, we also considered a scenario with clouds and haze removed. We computed the Bayes factor for this nested model and found a detection significance of 3.7σ for the cloud and haze parameters. The detection significance is higher than for the FREE model parameterization because the assumption of chemical equilibrium breaks the degeneracy between the cloud/haze and

the other molecular abundances (Na+K, TiO); i.e. the other molecular abundances can't increase arbitrarily to make up for the absence of clouds/haze.

5.3. Retrieval with NEMESIS

We also compared the results from CHIMERA with output from an independent retrieval code to test the robustness of our measurements against a different modeling approach (optimal estimation versus MCMC). We fit the high-resolution WFC3 spectrum with the NEMESIS code (Irwin et al. 2008). The model included H₂O as the only molecular opacity source. We modeled the atmosphere's thermal structure as a scalar multiple of the dayside temperature-pressure profile from Stevenson et al. (2014b). NEMESIS does not currently incorporate cloud-top pressure as a free parameter, so we fixed the altitude of an opaque gray cloud deck over a grid of pressures ranging from 10 to 10⁻³ mbar and retrieved the atmospheric properties for each case (for further discussion of retrieving cloud properties with NEMESIS, see Barstow et al. 2013). The best fit model had a cloud deck at 1 mbar and a retrieved water abundance of $4.0 \times 10^{-4} - 1.7 \times 10^{-3}$ at 1σ . Models with higher altitude cloud decks (0.1 - 0.001 mbar) achieved nearly as good a fit, and had 1σ upper bounds on the water abundance of 3×10^{-2} . These results are in excellent agreement with the results from CHIMERA.

6. IMPLICATIONS FOR ATMOSPHERIC C/O

In this section, we quantify the strength of the evidence for oxygen-rich compositions (C/O < 1) over carbon-rich compositions (C/O > 1) based on our retrieval results from two separate modeling approaches. We focus first on constraints from the water abundance from the FREE retrieval, since water is the only molecule that is unambiguously detected in the spectrum. We then discuss constraints from the C-C model, which retrieves C/O directly under the assumption of chemical equilibrium.

6.1. Constraints from the FREE Retrieval

Broadly speaking, a carbon-rich atmosphere is expected to have lower water abundance than an oxygen-rich atmosphere, because most of the oxygen atoms are bound in CO in chemical equilibrium at high temperatures (Madhusudhan et al. 2011a). Our water abundance measurement is a qualitatively better match to predictions for an oxygen-rich composition. In Figure 13, we show the retrieved H₂O abundance for the FULL retrieval and nested models in comparison with thermochemical equilibrium predictions for oxygen-rich (C/O = 0.55 = solar) and carbon-rich (C/O = 1) atmospheres. The predicted abundance depends on pressure and temperature, so we plot the span of predictions over pressures from 0.1 to 10 mbar as a function of temperature.

To quantitatively compare our measurement to the models, we must marginalize over the uncertainty in the pressure and temperature probed by the observations. To do this, we compute an equivalent pressure level p_{eq} corresponding to an optical depth $\tau = 0.56$ at 1.4 μm at each step in the MCMC chain. This quantity is representative of the typical pressure level at which photons are absorbed by water molecules (Lecavelier Des Etangs et al. 2008). Figure 14 shows the distribution of equivalent pressures from the MCMC chain. Note that p_{eq}

is not a retrieved quantity, but rather derived from the opacities at each MCMC step. We show the distribution of equivalent pressures obtained from this method in Figure 14. The pressures have a 1σ range of 0.05 to 5 mbar, with a peak at 0.5 mbar.

For a specified atmospheric composition, we can then calculate the predicted equilibrium H₂O abundance for a temperature and pressure equal to T_s and p_{eq} . This procedure yields a “calculated” water abundance to compare with the “observed” water abundance at each step in the MCMC chain. We show the distribution of observed minus calculated (“O - C”) values for the oxygen-rich and carbon-rich models in Figure 13 (right panel).

To test how well each model agrees with the retrieved H₂O abundance, we assess whether the O-C distribution is consistent with zero. For the oxygen-rich model, zero is contained in the 1σ credible interval centered on the median, indicating that this model is a good match. By contrast, the O-C distribution for the carbon-rich model is in tension with zero at approximately 2σ (the 95% credible interval is 0.1 – 6.4). The median O-C value is 3.7, implying that the typical retrieved water abundances are nearly four orders of magnitude larger than predicted for a carbon-rich composition.

We emphasize that these results are strongly dependent on the temperature and pressure probed by the observations. Assuming a different local thermal profile can result in order-of-magnitude differences in the predicted abundances (cf. Figure 2, Madhusudhan 2012). It is therefore essential to account for the temperature and pressure (and their uncertainties) when estimating C/O.

6.2. Constraints from the C-C Retrieval

We obtain more stringent constraints on the atmospheric C/O from the C-C retrieval than from the FREE parameterization. By every metric we enumerate below, C/O values greater than one are ruled out at high confidence. First, the entire posterior probability distribution for C/O from the fiducial model is less than one. We calculate that there are roughly 3000 independent samples in the MCMC chain for that model, which is a sufficient number to rule out C/O > 1 at greater than 3σ confidence. Second, the χ^2 values for the best fit models provide additional evidence in favor of the O-rich scenario. The best fit O-rich model has $\chi^2 = 2.75$ versus $\chi^2 = 29.2$ for the best fit C-rich case (for 2 degrees of freedom). This difference in χ^2 implies the O-rich model is 10^6 times more likely than the best fit C-rich model, assuming Gaussian statistics. Third, the Bayes factor (the ratio of integrated posterior probabilities) for the C-rich scenario to the O-rich scenario is 14, which constitutes strong evidence in favor of the O-rich model (Jeffreys 1998). In addition, the fact that the temperature range drops unphysically low – below the condensation temperature of water – is further evidence that the carbon-rich model is not appropriate for these data.

Taken together, these results rule out a carbon-rich composition for the atmosphere at high confidence. This is a more definitive constraint than what we obtained from studying the water abundance alone because the model is sensitive to both the presence of water *and* the absence of absorption features from other molecules. For example, the fact that no methane features are detected

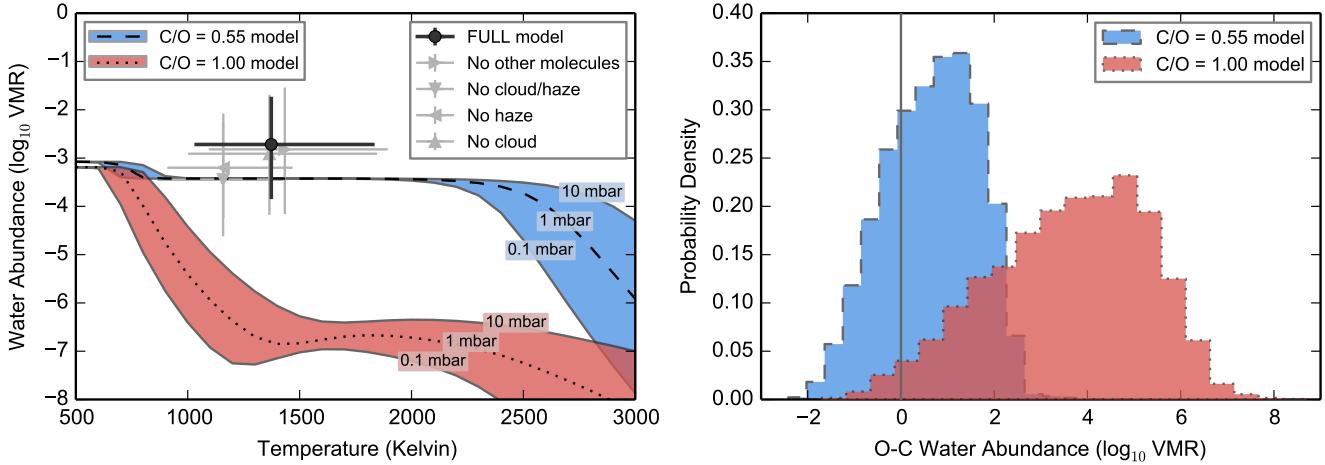


FIG. 13.— **Left:** Measurement of the water abundance and scale height temperature for WASP-12b (points) compared to equilibrium chemistry predictions of the water content for different atmospheric compositions (lines and shading). The black point indicates the water abundance and temperature measurements from the FULL model from the CHIMERA fit to the WFC3 spectrum. Results from other nested models are shown in gray. The black dashed line and blue shading correspond to water abundance predictions for a solar C/O composition, and the black dotted line and red shading correspond to C/O = 1. Both models have solar metallicity. For each model composition, the shading shows the span of predicted water abundances over pressures ranging from 0.1 - 10 mbar. The black lines correspond to 1 mbar, which is the typical pressure level probed by our observations. **Right:** Histogram of observed minus calculated water abundances for the FULL retrieval results relative to different model compositions. The red histogram (dotted line) shows the comparison with a carbon-rich model, and the blue histogram (dashed line) shows results for the oxygen-rich model. The gray vertical line marks zero, where the observations are an exact match to the model predictions.

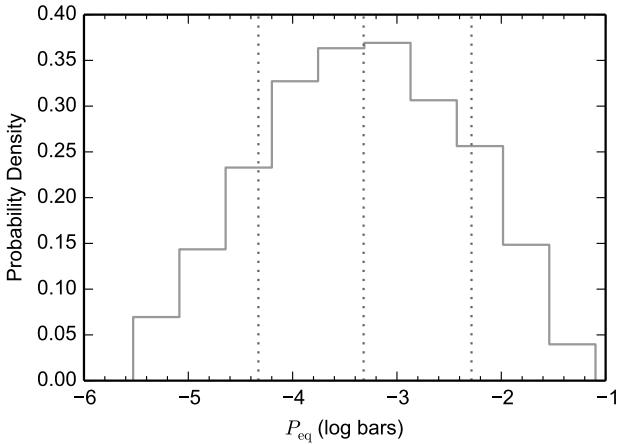


FIG. 14.— Histogram of equivalent pressures corresponding to an optical depth of 0.56 at $1.4 \mu\text{m}$ for each step in the MCMC chain for the FULL retrieval. The dotted lines indicate the median and surrounding 68% confidence interval.

in the spectrum strengthens the case for an O-rich composition beyond the constraints from the presence of water alone. The C-C model also assumes more prior knowledge by imposing chemical equilibrium in the calculation of the model spectra.

7. SUMMARY & CONCLUSIONS

We have measured a precise transmission spectrum for the hot Jupiter WASP-12b over the wavelength range 0.84 to $1.67 \mu\text{m}$ with *HST*/WFC3. The transmission spectrum is a factor of three more precise than previous measurements in this wavelength range (Swain et al. 2013). We retrieved the atmospheric properties based on this spectrum with a variety of models, including a new retrieval parameterization that fits for C/O and metallic-

ity rather than molecular abundances. Our conclusions about the nature of the planet are summarized as follows:

1. Water is present in the atmosphere. Models that do not include water absorption are excluded at 7σ confidence. This result is the first unambiguous spectroscopic detection of a molecule in the planet’s atmosphere.
2. The 68% credible interval for the retrieved water abundance is $10^{-5} - 10^{-2}$ for a wide range of models. This result applies to scenarios where water is the only absorber included in the model, where clouds and haze are added, where additional absorbers are incorporated (CO, CO₂, NH₃, TiO, VO, Na, K, CH₄, C₂H₂, HCN, H₂S, FeH, and N₂), and where optical transit depth measurements from *HST*/STIS are also fit. This range agrees well with the predicted equilibrium water abundance for a solar composition atmosphere but is still consistent to within 2σ with the prediction for a carbon-rich composition.
3. Stellar photons are absorbed at temperatures of 1000 – 1900 Kelvin and pressures of 0.1 – 10 mbar (based on 1σ ranges for T_s and P_{eq} from the FULL model). These estimates have large uncertainties because the transmission spectrum is only weakly sensitive to the thermal structure of the atmosphere.
4. Based on a new retrieval parameterization that fits for C/O and metallicity under the assumption of chemical equilibrium, we constrain the C/O to $0.5^{+0.2}_{-0.3}$ at 1σ and rule out a carbon-rich atmosphere composition (C/O > 1) at $> 3\sigma$ confidence. With this model, we also detect the presence of clouds/haze at 3.7σ confidence.

Our constraint on C/O is in tension with past studies of the planet's atmosphere that found a carbon-rich composition was the best explanation for the dayside emission spectrum (e.g. Madhusudhan et al. 2011b; Stevenson et al. 2014b). We would hope for better agreement between these analyses, as both results are based on very high precision data fit with state-of-the-art retrieval models. This methodology has yielded consistent results for other planets, notably the hot Jupiter WASP-43b, which shows excellent agreement between estimates of the composition from the dayside emission and transmission spectra (Kreidberg et al. 2014a).

We note that a caveat for our results from the C-C model is our assumption that the atmosphere is in chemical equilibrium. Moses et al. (2013) considered the effects of photochemistry and mixing for WASP-12b and found that the water abundance is either unchanged from equilibrium values (for solar composition), or pushed even lower (for a C/O = 1 composition). Based on these results, it is unlikely that our high observed water abundance is due to disequilibrium chemistry in a carbon-rich atmosphere. However, these calculations were for the planet's dayside and it would be worth exploring disequilibrium effects for the terminator region specifically in future work.

Another assumption in our model is that the temperature structure is well approximated as 1D. In reality, a photon's slant path through the atmosphere traverses many different temperatures and pressures. However, theoretical models predict that most of the stellar radiation is attenuated in a fairly localized region within a few degrees of the terminator (Fortney et al. 2010). The change in temperature and pressure over a region this size is small compared to the uncertainty in our estimates of the thermal profile, so our assumption of a 1D profile is unlikely to bias our assessment of the atmospheric composition. On the other hand, the observations integrate over the entire limb of the planet, and there could be large variations in temperature over this region due to atmospheric dynamics. It is therefore possible that fitting the limb-averaged spectrum with a single 1D thermal profile could bias the results.

One route to reconciling the disagreement between constraints on C/O from the transmission and emission spectra is to determine the planet's global thermal structure. This could be achieved by combining phase-resolved emission spectroscopy of the planet with 3D atmospheric circulation modeling, as has been done for WASP-43b (Stevenson et al. 2014d; Kataria et al. 2015). This combination would provide an independent measure of the temperature structure at the terminator to strengthen our interpretation of the transmission spectrum. It would also put the dayside temperature-pressure profile in context of the planet's global heat circulation and aid in understanding the dayside emission spectrum. In addition, a spectroscopic phase curve would allow us to estimate the water abundance for new regions of the planet's atmosphere. Furthermore, the spectroscopic phase curve amplitudes themselves could provide an additional diagnostic of the atmosphere composition. Day-night temperature differences for hot Jupiters are larger at lower pressures (Showman et al. 2009; Steven-

son et al. 2014d; Kataria et al. 2014, 2015), so light curves in absorption bands (which probe lower pressures) are expected to have larger amplitudes than those in spectral windows. Measuring the wavelength dependence of the phase curve amplitude could therefore provide an additional constraint on the atmospheric composition (Showman et al. 2009; Stevenson et al. 2014d; Kataria et al. 2014, 2015).

Finally, our results highlight the necessity of obtaining high-precision data with multiple observing techniques (transmission spectroscopy, dayside emission spectroscopy, and phase curves) in order to obtain robust constraints on the rich chemistry and physics of exoplanet atmospheres. Studying the atmosphere from more than one angle (literally) is key to providing a detailed understanding of its thermal structure and dynamics, which is needed to unambiguously determine the composition (and vice versa). In addition, high-precision measurements are essential for revealing the small features present in exoplanet spectra. The amplitude of these features has often been smaller than predicted for transiting planets, due to the presence of clouds or haze and shallower thermal profiles (e.g. Fortney et al. 2006; Charbonneau et al. 2002; Deming et al. 2013). Our WASP-12b spectrum is a new example of this for transmission measurements, with features crossing just two scale heights. We therefore advocate applying an intensive approach of high-precision spectroscopy from multiple angles to a larger sample of transiting exoplanets to shed light on their nature and origins. Such measurements will help develop the observing strategies needed for definitive characterization of exoplanet atmospheres, and prepare the community to make robust measurements of potentially habitable worlds with future facilities.

This work is based on observations made with the NASA/ESA Hubble Space Telescope that were obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. These observations are associated with program GO-13467. Support for this work was provided by NASA through a grant from the Space Telescope Science Institute, the National Science Foundation through a Graduate Research Fellowship (to L.K.), the Alfred P. Sloan Foundation through a Sloan Research Fellowship (to J.L.B.), the Packard Foundation through a Packard Fellowship for Science and Engineering (to J.L.B.), and the Sagan Fellowship Program (to K.B.S.) as supported by NASA and administered by the NASA Exoplanet Science Institute. G.W.H. and M.H.W. acknowledge support from Tennessee State University and the State of Tennessee through its Centers of Excellence program.

We thank Patrick Irwin for the use of his NEMESIS code. We are grateful to Megan Bedell, Hannah Diamond-Lowe, and Greg Gilbert for helpful discussions about this work. We thank Drake Deming, Peter McCullough, Adam Burrows, Sara Seager, Dave Charbonneau, and Derek Homeier for being co-investigators on the *HST* observing proposal.

REFERENCES

Ali-Dib, M., Mousis, O., Petit, J.-M., & Lunine, J. I. 2014, *ApJ*, 785, 125

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, 47, 481

Atreya, S. K., Mahaffy, P. R., Niemann, H. B., Wong, M. H., & Owen, T. C. 2003, *Planet. Space Sci.*, 51, 105

Atreya, S. K., Wong, M. H., Owen, T. C., et al. 1999, *Planet. Space Sci.*, 47, 1243

Barstow, J. K., Aigrain, S., Irwin, P. G. J., Fletcher, L. N., & Lee, J.-M. 2013, *MNRAS*, 434, 2616

Barstow, J. K., Aigrain, S., Irwin, P. G. J., et al. 2014, *ApJ*, 786, 154

Bechter, E. B., Crepp, J. R., Ngo, H., et al. 2014, *ApJ*, 788, 2

Benneke, B., & Seager, S. 2013, *ApJ*, 778, 153

Bergfors, C., Brandner, W., Daemgen, S., et al. 2013, *MNRAS*, 428, 182

Berta, Z. K., Charbonneau, D., Bean, J., et al. 2011, *ApJ*, 736, 12

Berta, Z. K., Charbonneau, D., Désert, J.-M., et al. 2012, *ApJ*, 747, 35

Brogi, M., de Kok, R. J., Birkby, J. L., Schwarz, H., & Snellen, I. A. G. 2014, *A&A*, 565, A124

Burrows, A., & Sharp, C. M. 1999, *ApJ*, 512, 843

Burton, J. R., Watson, C. A., Rodríguez-Gil, P., et al. 2015, *MNRAS*, 446, 1071

Campo, C. J., Harrington, J., Hardy, R. A., et al. 2011, *ApJ*, 727, 125

Castelli, F., & Kurucz, R. L. 2004, ArXiv Astrophysics e-prints, astro-ph/0405087

Charbonneau, D., Brown, T. M., Noyes, R. W., & Gilliland, R. L. 2002, *ApJ*, 568, 377

Copperwheat, C. M., Wheatley, P. J., Southworth, J., et al. 2013, *MNRAS*, 434, 661

Cornish, N. J., & Littenberg, T. B. 2007, *Phys. Rev. D*, 76, 083006

Cowan, N. B., Machalek, P., Croll, B., et al. 2012, *ApJ*, 747, 82

Croll, B., Lafreniere, D., Albert, L., et al. 2011, *AJ*, 141, 30

Croll, B., Albert, L., Jayawardhana, R., et al. 2015, *ApJ*, 802, 28

Crossfield, I. J. M., Barman, T., Hansen, B. M. S., Tanaka, I., & Kodama, T. 2012, *ApJ*, 760, 140

Deming, D., Wilkins, A., McCullough, P., et al. 2013, *ApJ*, 774, 95

Désert, J.-M., Sing, D., Vidal-Madjar, A., et al. 2011, *A&A*, 526, A12

Diamond-Lowe, H., Stevenson, K. B., Bean, J. L., Line, M. R., & Fortney, J. J. 2014, *ApJ*, 796, 66

Dodson-Robinson, S. E., Willacy, K., Bodenheimer, P., Turner, N. J., & Beichman, C. A. 2009, *Icarus*, 200, 672

Föhring, D., Dhillon, V. S., Madhusudhan, N., et al. 2013, *MNRAS*, 435, 2268

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306

Fortney, J. J., Cooper, C. S., Showman, A. P., Marley, M. S., & Freedman, R. S. 2006, *ApJ*, 652, 746

Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, *ApJ*, 678, 1419

Fortney, J. J., Shabram, M., Showman, A. P., et al. 2010, *ApJ*, 709, 1396

Fossati, L., Haswell, C. A., Froning, C. S., et al. 2010, *ApJ*, 714, L222

Gautier, D., Hersant, F., Mousis, O., & Lunine, J. I. 2001, *ApJ*, 550, L227

Guillot, T. 2010, *A&A*, 520, A27

Haswell, C. A., Fossati, L., Ayres, T., et al. 2012, *ApJ*, 760, 79

Hauschildt, P. H., Allard, F., & Baron, E. 1999, *ApJ*, 512, 377

Hebb, L., Collier-Cameron, A., Loeillet, B., et al. 2009, *ApJ*, 693, 1920

Heng, K., Hayek, W., Pont, F., & Sing, D. K. 2012, *MNRAS*, 420, 20

Hersant, F., Gautier, D., & Lunine, J. I. 2004, *Planet. Space Sci.*, 52, 623

Irwin, P. G. J., Teanby, N. A., de Kok, R., et al. 2008, *J. Quant. Spec. Radiat. Transf.*, 109, 1136

Jeffreys, H. 1998, *The Theory of Probability* (OUP Oxford)

Kass, R. E., & Raftery, A. E. 1995, *Journal of the american statistical association*, 90, 773

Kataria, T., Showman, A. P., Fortney, J. J., Marley, M. S., & Freedman, R. S. 2014, *ApJ*, 785, 92

Kataria, T., Showman, A. P., Fortney, J. J., et al. 2015, *ApJ*, 801, 86

Knutson, H. A., Charbonneau, D., Noyes, R. W., Brown, T. M., & Gilliland, R. L. 2007, *ApJ*, 655, 564

Knutson, H. A., Dragomir, D., Kreidberg, L., et al. 2014, *ApJ*, 794, 155

Konopacky, Q. M., Barman, T. S., Macintosh, B. A., & Marois, C. 2013, *Science*, 339, 1398

Kopparapu, R. k., Kasting, J. F., & Zahnle, K. J. 2012, *ApJ*, 745, 77

Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014a, *ApJ*, 793, L27

—. 2014b, *Nature*, 505, 69

Lecavelier Des Etangs, A., Pont, F., Vidal-Madjar, A., & Sing, D. 2008, *A&A*, 481, L83

Line, M. R., Knutson, H., Deming, D., Wilkins, A., & Desert, J.-M. 2013a, *ApJ*, 778, 183

Line, M. R., Knutson, H., Wolf, A. S., & Yung, Y. L. 2014, *ApJ*, 783, 70

Line, M. R., Vasish, G., Chen, P., Angerhausen, D., & Yung, Y. L. 2011, *ApJ*, 738, 32

Line, M. R., Wolf, A. S., Zhang, X., et al. 2013b, *ApJ*, 775, 137

Lodders, K. 2004, *ApJ*, 611, 587

López-Morales, M., Coughlin, J. L., Sing, D. K., et al. 2010, *ApJ*, 716, L36

Madhusudhan, N. 2012, *ApJ*, 758, 36

Madhusudhan, N., Amin, M. A., & Kennedy, G. M. 2014a, *ApJ*, 794, L12

Madhusudhan, N., Crouzet, N., McCullough, P. R., Deming, D., & Hedges, C. 2014b, *ApJ*, 791, L9

Madhusudhan, N., Mousis, O., Johnson, T. V., & Lunine, J. I. 2011a, *ApJ*, 743, 191

Madhusudhan, N., & Seager, S. 2009, *ApJ*, 707, 24

Madhusudhan, N., Harrington, J., Stevenson, K. B., et al. 2011b, *Nature*, 469, 64

Mandel, K., & Agol, E. 2002, *ApJ*, 580, L171

Mandell, A. M., Haynes, K., Sinukoff, E., et al. 2013, *ApJ*, 779, 128

McBride, B. J., & Gordon, S. 1996, Computer Program for Calculation of Complex Chemical Equilibrium Compositions and Applications II. User's Manual and Program Description, Tech. rep., NASA

McCullough, P., & MacKenty, J. 2012, Considerations for using Spatial Scans with WFC3, Tech. rep.

Moses, J. I., Madhusudhan, N., Visscher, C., & Freedman, R. S. 2013, *ApJ*, 763, 25

Moses, J. I., Visscher, C., Fortney, J. J., et al. 2011, *ApJ*, 737, 15

Nichols, J. D., Wynn, G. A., Goad, M., et al. 2015, ArXiv e-prints, arXiv:1502.07489

Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, *ApJ*, 743, L16

Owen, T., Mahaffy, P., Niemann, H. B., et al. 1999, *Nature*, 402, 269

Parmentier, V., & Guillot, T. 2014, *A&A*, 562, A133

Robinson, T. D., & Catling, D. C. 2012, *ApJ*, 757, 104

Sellke, T., Bayarri, M. J., & Berger, J. O. 2001, *The American Statistician*, 55, pp. 62

Showman, A. P., Fortney, J. J., Lian, Y., et al. 2009, *ApJ*, 699, 564

Sing, D. K., Lecavelier des Etangs, A., Fortney, J. J., et al. 2013, *MNRAS*, 436, 2956

Stevenson, K. B., Bean, J. L., Fabrycky, D., & Kreidberg, L. 2014a, *ApJ*, 796, 32

Stevenson, K. B., Bean, J. L., Madhusudhan, N., & Harrington, J. 2014b, *ApJ*, 791, 36

Stevenson, K. B., Bean, J. L., Seifahrt, A., et al. 2014c, *AJ*, 147, 161

Stevenson, K. B., Désert, J.-M., Line, M. R., et al. 2014d, *Science*, 346, 838

Swain, M., Deroo, P., Tinetti, G., et al. 2013, *Icarus*, 225, 432

Swain, M. R., Line, M. R., & Deroo, P. 2014, *ApJ*, 784, 133

Teske, J. K., Cunha, K., Smith, V. V., Schuler, S. C., & Griffith, C. A. 2014, *ApJ*, 788, 39

Teske, J. K., Schuler, S. C., Cunha, K., Smith, V. V., & Griffith, C. A. 2013, *ApJ*, 768, L12

Trotta, R. 2008, *Contemporary Physics*, 49, 71

Venot, O., Hébrard, E., Agúndez, M., Decin, L., & Bounaceur, R. 2015, ArXiv e-prints, arXiv:1502.03567

Weinberg, M. D. 2012, *Bayesian Anal.*, 7, 737

Zhao, M., Monnier, J. D., Swain, M. R., Barman, T., & Hinkley, S. 2012, *ApJ*, 744, 122